ON CHARACTERIZATIONS OF ORTHOGONAL DESIGNS

J. A. Eccleston

October, 1971

Abstract

Characterization of orthogonal designs is desirable and important, since under certain response models they can yield independent or uncorrelated estimators and "good" test statistics. This paper classifies orthogonal designs as either totally orthogonal or orthogonal with respect to certain parameters of the design. The concept of a totally orthogonal design is the same as the classical notion of an orthogonal design. Three approaches are utilized to derive definitions of orthogonal designs; they are combinatorial, covariance and difference of means and parameters. Some facts and properties of orthogonal designs are presented.

396
Number BU-~~393~~-M in the Biometrics Unit Mimeo Series, Cornell University.

ON CHARACTERIZATIONS OF ORTHOGONAL DESIGNS

J. A. Eccleston

BU-393-M October 1971

1. Introduction

The ease with which a statistical analysis can be performed on a set of data depends largely on the experimental design or layout used to collect the data. If an experiment is well designed then the computations necessary to test hypotheses and/or constructing confidence intervals are quite simple; even complex experiments can be handled by repeated applications of essentially one process. This simplicity and well designed structure of the layout depend to a large extent on whether or not the design is orthogonal. Orthogonality as described by Yates [1] is that property of design which ensures different classes of effects (parameters) direct and separate estimation without entanglement. By "without entanglement", Yates is implying that the estimators of different classes of effects be unrelated (i.e., zero covariance).

Characterization of orthogonal designs is both desirable and necessary since they can, under certain response models, yield independent or uncorrelated estimators and good test statistics (see section 4,F-6) enabling accurate probability statements and inferences to be made from the data. In the following sections several definitions of orthogonal designs are discussed with examples demonstrating situations when definitions agree or disagree as to the orthogonality of a particular design. Some facts and properties concerning orthogonal designs are also presented.

Biometrics Unit, Cornell University, Ithaca, New York.

2. Some Definitions of Orthogonality

2.1. Covariance definition.

Consider a design D with the response model $\underline{y} = X\underline{\psi} + \underline{\epsilon}$ where y is the vector of observed responses, X is the design matrix, and $\underline{\psi} = (C_1, C_2, \cdots, C_m)$, C_i is a set of parameters with cardinality c_i, $i=1,2,\cdots,m$. Then D is said to be orthogonal with respect to the sets C_i and C_j, $i \neq j$, if under the given response model

$$\mathrm{Cov}\left(\hat{\theta}^i_\ell, \hat{\theta}^j_k\right) = 0 \;, \quad \theta^i_\ell \in C_\ell \text{ and } \theta^j_k \in C_j$$

$$i \neq j,\ \ell = 1,2,\cdots,c_i \;,\ k=1,2,\cdots,c_j \;.$$

($\hat{\theta}^i_\ell$ and $\hat{\theta}^j_k$ are estimators of θ^i_ℓ and θ^j_k respectively.) D is said to be <u>totally orthogonal</u> if the covariance between estimators of any two parameters belonging to any two arbitrary, but different, sets of parameters (as defined above) is zero.

Example 2.1. Consider a (3,2,1) Youden design

Columns (b_k)

rows (r_j)

1	2	3
2	3	1

treatments coded as 1, 2, 3

$$\text{Response model: } y_{ijk} = \mu + t_i + r_j + b_k + \epsilon_{ijk}$$

where μ = mean effect, t_i = effect of i^{th} treatment, r_j = effect of j^{th} row and b_k = effect of k^{th} column. $i=1,2,3$, $j=1,2$, $k=1,2,3$. ϵ_{ijk}'s are independently distributed with zero mean and variance σ^2, for all i j and k.

By the Theory of Least Squares and constraints $\sum_{i=1}^{3} t_i = 0$, $\sum_{j=1}^{2} r_j = 0$ and $\sum_{k=1}^{3} b_k = 0$ we obtain estimators for μ, t_i, r_j and c_k

$$\hat{\mu} = \bar{y}_{...}$$

$$\hat{t}_1 = (y_{111}+y_{123})/2 - \hat{\mu} + \hat{b}_2/2$$

$$\hat{t}_2 = (y_{221}+y_{212})/2 - \hat{\mu} + \hat{b}_3/2$$

$$\hat{t}_3 = (y_{322}+y_{313})/2 - \hat{\mu} + \hat{b}_1/2$$

$$\hat{b}_1 = (y_{111}+y_{221})/2 - \hat{\mu} + \hat{t}_3/2$$

$$\hat{b}_2 = (y_{212}+y_{322})/2 - \hat{\mu} + \hat{t}_1/2$$

$$\hat{b}_3 = (y_{313}+y_{123})/2 - \hat{\mu} + \hat{t}_2/2$$

$$\hat{r}_1 = \bar{y}_{.1.} - \hat{\mu}$$

$$\hat{r}_2 = \bar{y}_{.2.} - \hat{\mu}$$

where

$$\bar{y}_{...} = \sum_{i=1}^{3} \sum_{j=1}^{2} \sum_{k=1}^{3} \frac{y_{ijk}}{6} \text{ and } \bar{y}_{.j.} = \sum_{i=1}^{3} \sum_{k=1}^{3} \frac{y_{ijk}}{3} .$$

It is easily shown that

$$\text{Cov}(\hat{t}_i \hat{r}_j) = 0 \quad \text{for all } i \text{ and } j$$

and

$$\text{Cov}(\hat{r}_j \hat{b}_k) = 0 \quad \text{for all } j \text{ and } k$$

but

$$\text{Cov}(\hat{t}_i \hat{b}_k) \neq 0 \quad \text{for all } i \text{ and } k .$$

Thus it follows that rows and columns are orthogonal, rows and treatments are orthogonal, but treatments and columns are not. If an appropriate third row is added to the design, namely 3 in column 1, 1 in column 2, and 2 in column 3 then the resulting design is a latin square of order 3 and is totally orthogonal.

A paper by Darroch and Silvey [2] and one by Seber [3] define orthogonality from the point of view of linear hypotheses, concerning parameters of the design, partitioning the sample space into perpendicular subspaces.

2.2. "Nested hypotheses" definition.

Let $L: \underline{y} = \underline{\theta} + \underline{\epsilon}$ be the general linear model where $\underline{\epsilon} \sim N(0, \sigma^2 I)$ and $\underline{\theta}$ is the vector of parameters belonging to the sampling space, Ω, a subspace on the n-dimensional Euclidean space R^n. Consider a sequence of nested linear hypotheses $H_i: \underline{\theta} \in w_i$, $i=1,\cdots,k$ where w_i is a subspace of Ω. Then from Darroch and Silvey [2] and Seber [3] an experimental design is orthogonal relative to the general linear model L and the nested linear hypotheses $H_1, H_2, \cdots, H_k$ if the subspaces $w_1, w_2, \cdots, w_k$ are such that $w_i^{\perp} \cap \Omega \perp w_j^{\perp} \cap \Omega$ for all i and j, $i \neq j$ (i.e., if the intersection of the orthogonal complements of w_i and w_j, $i \neq j$, $i,j = 1,2,\cdots,k$ with Ω are mutually perpendicular).

If L is the analysis of variance model then the sums of squares obtained by nesting the hypotheses are stochastically independent and are the same irrespective of the order of nesting. (See Scheffé [4].)

Remark. It is clear that the "hypotheses" definition expresses in general terms the covariance definition (2.1).

Lemma 2.1. *If* $\Omega = \{\theta | A\theta = 0\}$, $w_i = \{\theta | A\theta = 0,\ A_i\theta = 0\}$ *where the rows of the matrix* $[A' \vdots A_i']'$ *are linearly independent* $(i=1,2,\cdots,k)$ *and* $AA_i' = 0$ for $i=1,2,\cdots,k$ *then* $w_i^{\perp} \cap \Omega \perp w_j^{\perp} \cap \Omega$ *if and only if* $A_iA_j' = 0$. (For proof see Seber [3].)

An interesting and somewhat more geometrical "nested hypotheses" definition is given by Scheffé [4] as follows: Consider the linear model L and nested linear hypotheses with $w_i = \Omega \cap H_1 \cap H_2 \cap \cdots \cap H_i$ ($i=1 \cdots k$) as before. Let V_i denote the space in which $\underline{\theta}$ is constrained to lie, under w_i so that $V_1 \supset V_2 \supset \cdots \supset V_k$ and also let $\underline{\theta}_{w_i}$ denote the projection of $\underline{y}$ on V_i. The vector $\underline{y}$ can now be partitioned into a sum of k+2 mutually orthogonal vectors,

$$\underline{y} = (\underline{y}-\hat{\underline{\theta}}) + (\hat{\underline{\theta}}-\hat{\underline{\theta}}_{w_1}) + \cdots + (\hat{\underline{\theta}}_{w_{k-1}}-\hat{\underline{\theta}}_{w_k}) + \hat{\underline{\theta}}_{w_k}$$

and thus

$$\|\underline{y}\|^2 = \|y-\hat{\underline{\theta}}\|^2 + \|\hat{\underline{\theta}}-\hat{\underline{\theta}}_{w_1}\|^2 + \cdots + \|\hat{\underline{\theta}}_{w_{k-1}}-\hat{\underline{\theta}}_{w_k}\|^2 + \|\hat{\underline{\theta}}_{w_k}\|^2 \tag{2.1}$$

which in terms of sums of squares is

$$\text{S.S. total} = \text{S.S. error} + \text{S.S.}w_1 + \cdots + \text{S.S.}w_k + \text{S.S."mean"} .$$

If the projections of $\underline{y}$ on the V_i's are mutually orthogonal then the sums of squares will be independent and, with the normality assumption, distributed as independent chi-squares. As an example consider $L: y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$, $i=1 \cdots I$, $j=1 \cdots J$, $k=1 \cdots K$, ϵ_{ijk} are independently $N(0,\sigma^2)$. Now $\Omega = (\mu + \alpha_i + \beta_j + \gamma_{ij})$

$$H_1 : \gamma_{ij} = 0 \Rightarrow w_1 = (\mu+\alpha_i+\beta_j)$$

$$H_2 : \beta_j = 0 \Rightarrow w_2 = (\mu+\alpha_i)$$

$$H_3 : \alpha_i = 0 \Rightarrow w_3 = (\mu)$$

It follows that $\hat{\underline{\theta}} = [\bar{y}_{ij.}]$, $IJ \times 1$ vector

$$\underline{\hat{\theta}} - \underline{\hat{\theta}}_{w_1} = [\hat{\gamma}_{ij}] = [\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}-\bar{y}_{...}] \text{ , } IJ \times 1 \text{ vector}$$

$$\underline{\hat{\theta}}_{w_1} - \underline{\hat{\theta}}_{w_2} = [\hat{\beta}_j] = [\bar{y}_{.j.}-\bar{y}_{...}] \text{ , } J \times 1 \text{ vector}$$

$$\underline{\hat{\theta}}_{w_2} - \underline{\hat{\theta}}_{w_3} = [\hat{\alpha}_i] = [\bar{y}_{i..}-\bar{y}_{...}] \text{ , } I \times 1 \text{ vector}$$

and

$$\underline{\hat{\theta}}_{w_3} = \hat{\mu} = [\bar{y}_{...}]$$

where

$$\bar{y}_{ij.} = \sum_{k=1}^{K} y_{ijk}/K \text{ .}$$

Then equation 2.1, when written in terms of sums, i.e., $\|y\|^2 = \sum_{ijk} y^2_{ijk}$, is

$$\sum_i\sum_j\sum_k y^2_{ijk} = \sum_i\sum_j\sum_k (y_{ijk}-\bar{y}_{ij.})^2 + K\sum_i\sum_j (\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}-\bar{y}_{...})^2$$

$$+ IK\sum_j (\bar{y}_{.j.}-\bar{y}_{...})^2 + JK\sum_i (\bar{y}_{i..}-\bar{y}_{...})^2 + IJK\,\bar{y}^2_{...}$$

i.e., S.S. total = S.S.(error) + S.S.$(\hat{\gamma}_{ij})$ + S.S.$(\hat{\beta}_j)$ + S.S.$(\hat{\alpha}_i)$ + S.S.(mean) .

The sums of squares are independent and chi-square distributed.

2.3. Combinatorial definition.

Let D be a design with classes of effects (parameters) $c_{i_1}, c_{i_2}, \cdots, c_{i_p}$, $i_j = 1,\cdots,I_j$, and define $N(i_\ell i_k)$ to be a matrix of marginal frequencies of the classes c_{i_ℓ} and c_{i_k}. Then

$$N(i_\ell i_k) = [n_{ij}(i_\ell i_k)]$$

where

$n_{ij}(i_\ell i_k)$ = number of experimental units receiving the i^{th} level of c_{i_ℓ} and the j^{th} level of c_{i_k} .

$N(i_\ell i_k)$ is essentially the incidence matrix of c_{i_ℓ} and c_{i_k}, i.e.,

$$N(i_\ell i_k) = \begin{bmatrix} n_{11}(i_\ell i_k) & n_{12}(i_\ell i_k) & \cdots & n_{1I_k}(i_\ell i_k) \\ n_{21}(i_\ell i_k) & n_{22}(i_\ell i_k) & \cdots & n_{2I_k}(i_\ell i_k) \\ \vdots & \vdots & \ddots & \vdots \\ n_{I_\ell 1}(i_\ell i_k) & n_{I_\ell 2}(i_\ell i_k) & \cdots & n_{I_\ell I_k}(i_\ell i_k) \end{bmatrix}$$

$$n_i(i_\ell) = \sum_{j=1}^{I_k} n_{ij}(i_\ell i_k) \ , \quad n_j(i_k) = \sum_{i=1}^{I_\ell} n_{ij}(i_\ell i_k) \text{ and } n = \sum_i^{I_\ell} \sum_j^{I_k} n_{ij}(i_\ell i_k)$$

$n_i(i_\ell)$ = number of experimental units receiving i level of c_{i_ℓ} .

[NOTE: In matrix notation, consider $\underline{1}$ as the row vector of ones, then $1N(i_\ell i_k)$ = vector of column sums, $N(i_\ell i_k)1'$ = vector row sums and $1N1' = n$ = total number of observations.]

D is said to be orthogonal with respect to classes c_{i_ℓ} and c_{i_k} if all elements of $N(i_\ell i_k)$ are such that

$$n_{ij}(i_\ell i_k) = \frac{n_i(i_\ell) n_j(i_k)}{n} \quad \text{for all } i \text{ and } j \ . \tag{2.2}$$

If equation (2.2) is true for all pairs of classes c_{i_ℓ} and c_{i_k} $\ell \neq k$ $\ell,k=1,2,\cdots,p$ then D is said to be <u>totally</u> orthogonal.

A combinatorial definition more generally used is the following. The design D with classes $c_{i_1}, c_{i_2}, \cdots, c_{i_p}$ is said to be (totally) orthogonal if

$$n_{i_1',i_2',\cdots,i_p'}(i_1\, i_2 \cdots i_p) = \frac{n_{i_1'}(i_1)\, n_{i_2'}(i_2) \cdots n_{i_p'}(i_p)}{(n)^{p-1}}$$

for all $i_1'\ i_2' \cdots i_p'$

$n_{i_1',i_2',\cdots,i_p'}(i_1\, i_2 \cdots i_p)$ = number of experimental units receiving the i_1' level of c_{i_1}, i_2' level of c_{i_2}, $\cdots$, and the i_p' level of c_{i_p}

with $n_{i_1'}(i_1),\cdots,n_{i_p'}(i_p)$ and n as before.

Example 2.2. Consider the (3,2,1) Youden design given in the example of Section 2.1.

$$N(ij) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \text{treatment-row marginal matrix}$$

$$n_1(i) = 2,\quad n_2(i) = 2,\quad n_3(i) = 2,\quad n_1(j) = 3,\quad n_2(j) = 3, \text{ and } n = 6 .$$

From equation (2.2) we get

$$\begin{array}{lll} n_{11}(i,j) = 1 & n_{21}(i,j) = 1 & n_{31}(i,j) = 1 \\ n_{12}(i,j) = 1 & n_{22}(i,j) = 1 & n_{32}(i,j) = 1 \end{array}$$

Thus treatments and rows are orthogonal ($t_i \perp r_j$) and similarly rows and columns are orthogonal, but treatments and columns are not orthogonal, as the following shows:

$$N(ik) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$n_1(i) = n_2(i) = n_3(i) = 2 \quad \text{and} \quad n_1(k) = n_2(k) = n_3(k) = 2$$

Thus

$$n_{11}(ik) = \frac{2\times 2}{6} \neq 1 \quad .$$

Example 2.3. Consider a 3-way model with interactions and unequal numbers.

$$\Omega : y_{ijk\ell} = \mu + \alpha_i + \beta_j + \delta_k + (\alpha\beta)_{ij} + (\alpha\delta)_{ik} + (\beta\delta)_{jk} + (\alpha\beta\delta)_{ijk} + \epsilon_{ijk\ell}$$

$$i=1 \cdots I, \; j=1 \cdots J, \quad k=1 \cdots K, \quad \ell=1 \cdots L_{ijk}$$

$$n_{111}(ijk) = L_{111}, \quad n_{112}(ijk) = L_{112}, \; \cdot \; \cdot \; \cdot, \; n_{IJK}(ijk) = L_{IJK} \quad .$$

To test whether α_i β_j and δ_k are orthogonal to each other is just a matter of constructing the appropriate $N(\cdot\cdot)$s and using equation (2.2). The definition also applies to interaction terms and μ.

$$N[i,(i,j)] = \begin{bmatrix} n_{1,11}(i,ij) & n_{1,12}(i,ij) & \cdots & n_{1,1J}(i,ij) \\ n_{2,21}(i,ij) & n_{2,22}(i,ij) & \cdots & n_{2,2J}(i,ij) \\ \vdots & \vdots & \ddots & \vdots \\ n_{I,I1}(i,ij) & n_{I,I2}(i,ij) & \cdots & n_{I,IJ}(i,ij) \end{bmatrix}$$

$n_{m,mn}[i,(ij)]$ = number of experimental units receiving the m level of α_i and the mn level of $(\alpha\beta)_{ij}$

or $n_{m,mn}[i,(ij)]$ = number of experimental units receiving the mn level of $(\alpha\beta)_{ij}$.

For $(\alpha\beta)_{ij}$ and $(\beta\delta)_{jk}$

$n_{mn,mp}[(ij)(jk)]$ = number of experimental units receiving the mn level of $(\alpha\beta)_{ij}$ and the np level of $(\beta\delta)_{jk}$

or $n_{mn,np}[(ij)(jk)]$ = number of experimental units receiving the m level of α_i, n level of β_j, and the p level of δ_k .

For μ and $(\alpha\beta)_{ij}$, say

$n_{mm}[(ij)]$ = number of experimental units receiving mn levels of $(\alpha\beta)_{ij}$ and μ,

and so on for the remaining marginal matrices.

2.4. Difference of means definition.

Consider a design D with, say, p sets of parameters or classes of effects $c_{i_1}, c_{i_2}, \cdots, c_{i_p}$ and responses $y_{i_1,i_2,\cdots,i_p}$, $i_1=1 \cdots I_1, i_2=1 \cdots I_2, \cdots, i_p=1 \cdots I_p$. If the differences between the means of the responses at two different levels of a parameter involve the difference of the levels of that parameter, a function of the random error and no other parameter(s) in any way then the parameter in question is orthogonal to all other parameters of the design. More simply, if the response model is

$$\underline{y} = X\underline{c} + \underline{\epsilon}$$

$\underline{y}$ vector of responses, X is the design matrix, $\underline{c}$ is the vector of parameters and $\underline{\epsilon}$ is the vector of random errors. Then

$$\bar{y}_{\cdots i_j \cdots} - \bar{y}_{\cdots i_j' \cdots} = c_{i_j} - c_{i_j'} + f(\epsilon) \tag{2.3}$$

for all values of i_j, $i_j \neq i_j'$

implies that the c_{i_j} class is orthogonal to all other parameters in the model for all values of $i_j = 1 \cdots I_j$.

If equation (2.3) involves parameters other than c_{i_j} and $c_{i_j'}$ then those parameters and c_{i_j}'s are not orthogonal. More directly, any parameters appearing in an equation of the form of equation (2.3) are not orthogonal.

i.e. $$\bar{y}_{\ldots i_j \ldots} - \bar{y}_{\ldots i_j' \ldots} = c_{i_j} - c_{i_j'} + f(c_{i_\ell}\, c_{i_k}\, \epsilon)$$

for any value of i_j , $i_j \neq i_j'$.

Then c_{i_j} is not orthogonal to c_{i_ℓ} and c_{i_k}, but is orthogonal to all other parameters, namely c_{i_m}, for all m, $m \neq j$, $m \neq \ell$, $m \neq k$.

<u>Example 2.4</u>. Again consider the (3,2,1) Youden design previously used as an example. It is easy to show that

$$\bar{y}_{1..} - \bar{y}_{2..} = t_1 - t_2 - c_2 + c_3 + \epsilon_{111} + \epsilon_{123} - \epsilon_{221} - \epsilon_{212}$$

$$\bar{y}_{1..} - \bar{y}_{3..} = t_1 - t_3 - c_2 + c_1 + \epsilon_{111} + \epsilon_{123} - \epsilon_{322} - \epsilon_{312}$$

$$\bar{y}_{2..} - \bar{y}_{3..} = t_2 - t_3 - c_3 + c_1 + \epsilon_{221} + \epsilon_{212} - \epsilon_{322} - \epsilon_{312} .$$

A similar set of equations is derived for $\bar{y}_{..k} - \bar{y}_{..k'}$ (means of the columns). For rows we have

$$y_{.1.} - y_{.2.} = r_1 - r_2 + \bar{\epsilon}_{.1.} - \bar{\epsilon}_{.2.} \quad .$$

Thus treatments and columns are not orthogonal, but rows and columns, and rows and treatments are orthogonal.

3. Equivalence of Definitions

The obvious question now is, when and under what circumstances are the definitions of the previous section equivalent. By the use of a few rather simple examples it is shown that equivalence is not true in general.

Example 3.1. Consider a design D with parameters α_i, i=1,2, and β_j, j=1,2 and response model:

$$\Omega : \begin{cases} y_{ij} = \alpha_i\beta_j + \epsilon_{ij} \\ i=1,2, \quad j=1,2 \\ \epsilon_{ij}\text{'s independnetly } N(0,\sigma^2) \end{cases}$$

Let there be one observation at each level of α_i and β_j, i.e., 4 observations in all. The incidence matrix of α_i and β_j is

$$N(ij) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

which satisfies the combinatorial definition of orthogonality of α_i and β_j.

By the least squares method we obtain estimators of α_i and β_j

$$\hat{\alpha}_i = \frac{\sum_{j=1}^{2} y_{ij}\hat{\beta}_j}{\sum_j \hat{\beta}_j^2} \qquad \text{and} \qquad \hat{\beta}_j = \frac{\sum_{i=1}^{2} y_{ij}\hat{\alpha}_i}{\sum_i \hat{\alpha}_i^2}$$

Obviously it follows that

$$\text{Cov}(\hat{\alpha}_i\hat{\beta}_j) \neq 0 \quad .$$

Thus α_i and β_j are not orthogonal by the covariance definition.

Now consider the difference of means definition; we have

$$\bar{y}_{1.} - \bar{y}_{2.} = (\alpha_1-\alpha_2) \sum_{j=1}^{2} \beta_j + \bar{\epsilon}_{1.} - \bar{\epsilon}_{2.} \quad .$$

If the typical constraints $\sum_j \beta_j = 0$ and $\sum_i \alpha_i = 0$ are assumed then $\bar{y}_{1.} - \bar{y}_{2.} = \bar{\epsilon}_{1.} - \bar{\epsilon}_{2.}$. In both equations the definition does not hold. Thus α_i and β_j are not orthogonal by the difference of means definition.

<u>Example 3.2</u>. Let D be a design with parameters α_i and β_j and response model:

$$\Omega : \begin{cases} y_{ij} = \alpha_i + \beta_j + \alpha_i\beta_j + \epsilon_{ij} \\ i = 1 \cdots I \qquad j = 1 \cdots J \\ \epsilon_{ij}\text{'s independently } N(0,\sigma^2), \ \sum_i \alpha_i = 0, \ \sum_j \beta_j = 0 \end{cases}$$

The combinatorial definition is not considered in this example, since the "non-equivalence" of it and the other two definitions has been demonstrated in example 3.1.

By the method of least squares we obtain the following estimators for α_i and β_j:

$$\hat{\alpha}_i = \frac{\sum_j y_{ij}(1+\hat{\beta}_j) - \sum_j \hat{\beta}_j^2}{J + \sum_j \hat{\beta}_j^2}$$

and

$$\hat{\beta}_j = \frac{\sum_i y_{ij}(1+\hat{\alpha}_i) - \sum_i \hat{\alpha}_i^2}{I + \sum_i \hat{\alpha}_i^2} \quad .$$

Clearly the covariance of $\hat{\alpha}_i$ and $\hat{\beta}_j$ is not zero. Thus by the covariance definition α_i and β_j are not orthogonal.

But the difference of means definition gives

$$\bar{y}_{i.} - \bar{y}_{i'.} = \alpha_i - \alpha_{i'} + \bar{\epsilon}_{i.} - \bar{\epsilon}_{i'.}$$

and

$$\bar{y}_{.j} - \bar{y}_{.j'} = \beta_j - \beta_{j'} + \bar{\epsilon}_{.j} - \bar{\epsilon}_{.j'}$$

which means that α_i and β_j are orthogonal by the difference of means definition.

If the above design is totally crossed, that is if each level of i appears with each level of j and with an equal number of observations at each level of i and j then the combinatorial definition of orthogonality will hold.

From these examples we can state the following theorem.

Theorem 3.1. *The three definitions of orthogonality, the combinatorial, the covariance and the difference of means definitions, are in general not equivalent.*

It is interesting to notice that both of the above examples concern designs with response models that are nonlinear. If the class of designs under study is restricted to designs that have response models that are linear models (as defined by Graybill [5]) then the definitions are equivalent.

Theorem 3.2. *For the class of designs that have linear models as response models the definitions of orthogonality, the covariance, combinatorial and difference of means, are equivalent.*

Proof. Consider the general linear model

$$y_{i_1 i_2 \cdots i_{p+1}} = c_{i_1} + c_{i_2} + \cdots + c_{i_p} + \epsilon_{i_1 i_2 \cdots i_{p+1}}$$

where $i_j = 1,2,\cdots,I_j$, $j = 1,\cdots,p+1$, the c_{i_j}'s are classes of parameters and $\epsilon_{i_1 i_2 \cdots i_{p+1}}$ independently $N(0,\sigma^2)$.

Let the response $y_{i_1 i_2 \cdots i_{p+1}}$ be observed $k_{i_1 i_2 \cdots i_p}$ times and the usual constraints on the parameters be applied, i.e., $\sum_{i_j} k_{i_j} c_{i_j} = 0$. (It should be noted that some of the c_{i_j}'s may be interaction terms.)

Part (i): Equivalence of Combinatorial and Difference of Means Definitions.

From the model we have:

$$\bar{y}_{\cdots i_j \cdots} - \bar{y}_{\cdots i_j' \cdots} = c_{i_j} - c_{i_j'} + \sum_{i_1}\left(\frac{k_{i_1 i_j} c_{i_1}}{k_{i_j}} - \frac{k_{i_1 i_j'} c_{i_1}}{k_{i_j'}}\right) + \cdots$$

$$+ \sum_{i_p}\left(\frac{k_{i_j i_p} c_{i_p}}{k_{i_j}} - \frac{k_{i_j' i_p} c_{i_p}}{k_{i_j'}}\right) + \bar{\epsilon}_{\cdots i_j \cdots} - \bar{\epsilon}_{\cdots i_j' \cdots} \tag{3.2}$$

where $i_j \neq i_{j'}$, and

$$k_{i_j i_m} = \underset{(\text{not } i_j,\ \text{not } i_m)}{\sum_{i_1}^{I_1} \sum_{i_2}^{I_2} \cdots \sum_{i_p}} k_{i_1 i_2 i_3 \cdots i_p}$$

$$k_{i_j} = \underset{(\text{not } i_j)}{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_p}} k_{i_1 i_2 \cdots i_p} \quad \text{and} \quad k = \sum_{i_1} \sum_{i_2} \cdots \sum_{i_p} k_{i_1 i_2 \cdots i_p}$$

If c_{i_j} is orthogonal to c_{i_m} then c_{i_m} will not appear in equation (3.2), which implies that

$$\sum_{i_m}\left(\frac{k_{i_j i_m} c_{i_m}}{k_{i_j}} - \frac{k_{i_j' i_m} c_{i_m}}{k_{i_j'}}\right) = 0 \quad . \tag{3.3}$$

The constraint applied to c_{i_m} is $\sum_{i_m} k_{i_m} c_{i_m} = 0$. Equation (3.3) implies either or both of the following:

(i) $$\sum_{i_m} \frac{k_{i_j i_m} c_{i_m}}{k_{i_j}} = 0$$

and/or (ii) if we can write (3.3) as

$$\sum_{i_m}\left(\frac{k_{i_j i_m}}{k_{i_j}} - \frac{k_{i_j' i_m}}{k_{i_j'}}\right) c_{i_m} = 0$$

(i) gives

$$\sum_{i_m} \frac{k_{i_j i_m} c_{i_m}}{k_{i_j}} = \sum_{i_m} k_{i_m} c_{i_m} \quad \text{for all values of } i_j \text{ and } i_m$$

Thus

$$k_{i_j i_m} = k_{i_j} k_{i_m} \quad .$$

Summing over i_j and i_m we get k=1 (i.e, only one observed in the experiment) which is absurd.

(ii) implies $$\frac{k_{i_j i_m}}{k_{i_j}} - \frac{k_{i_j' i_m}}{k_{i_j'}} = 0 \quad \text{for values of } i_j \text{ and } i_m \; .$$

Summing over i_j' ($i_j' \neq i_j$) we have

$$k_{i_j i_m} = \frac{k_{i_j} k_{i_m}}{k} \quad . \tag{3.6}$$

Thus the difference of means definition implies equation (3.6). Now construct the marginal matrix for c_{i_j} and c_{i_m}:

$$N(i_j i_m) = [k_{i_j i_m}]$$

with row totals k_{i_j} and column totals k_{i_m}. Equation (3.6) is precisely the combinatorial relationship required for the combinatorial definition to hold. Similarly by arguing "backwards" the combinatorial definition gives rise to the difference of means definition.

Part (ii). Equivalence of Covariance and Difference of Means Definitions.

By the least squares method of estimation we have

$$\hat{c}_{i_j} = \bar{y}_{\dots i_j \dots} - \sum_{i_1} \frac{k_{i_1 i_j}}{k_{i_j}} \hat{c}_{i_1} - \cdots - \sum_{i_p} \frac{k_{i_j i_p}}{k_{i_j}} \hat{c}_{i_p} \tag{3.7}$$

when the difference of means definition (or equivalently, the combinatorial definition) holds, then we know that $k_{i_j i_m} = (k_{i_j} k_{i_m})/k$ is true for all values of i_j and i_m. Then it follows that equation (3.7) does not involve $\hat{c}_{i_m}$ in any way at all; thus the covariance of $\hat{c}_{i_j}$ and $\hat{c}_{i_m}$ is zero, also $\hat{c}_{i_m}$ has zero covariance with all other $\hat{c}_{i_\ell}$'s, $\ell \neq j$ and $\ell \neq m$, that appear in equation (3.7). Now if

$$\mathrm{Cov}(\hat{c}_{i_j} \hat{c}_{i_m}) = 0$$

then

$$\mathrm{Cov}(\bar{y}_{\dots i_j \dots} \; \hat{c}_{i_m}) = 0 \quad .$$

Thus

$$\mathrm{Cov}[(\bar{y}_{\dots i_j \dots} - \bar{y}_{\dots i_j' \dots}), \hat{c}_{i_m}] = 0$$

and $\hat{c}_{i_m} = c_{i_m} + f(\epsilon)$. Therefore $\bar{y}_{\ldots i_j \ldots} - \bar{y}_{\ldots i_j' \ldots} = c_{i_j} - c_{i_j'} + f(c_{i_\ell}\text{'s } \epsilon)_{\ell \neq m}$ not involving c_{i_m} in any way. Thus the difference of means definition is true. We now have

Covariance definition $\Leftrightarrow$ Difference of means definition (Part ii)

Combinatorial definition $\Leftrightarrow$ Difference of means definition (Part I)

Thus it follows that

Combinatorial definition $\Leftrightarrow$ Covariance definition.

That is, the definitions are equivalent. This completes the proof of theorem (3.2).

4. Some Facts

This section contains some trivial and perhaps not so trivial facts concerning orthogonality.

F.1. A parameter is not orthogonal to itself.

F.2. If a and b are parameters and a is orthogonal to b, $a \perp b$, then b is orthogonal to a, $b \perp a$ (i.e., orthogonality has the symmetry property).

F.3. If a, b and c are parameters, such that $a \perp b$ and $b \perp c$ then a and c are not necessarily orthogonal, i.e., orthogonality does not have the transitivity property. For example, consider the (3,2,1) Youden design used as an example in section 2.1. The design is such that treatments $\perp$ rows and rows $\perp$ columns but treatments are not orthogonal to columns.

F.4. If a_i, b_j and c_k are parameters of a linear model such that a_i is not orthogonal to b_j and b_j is not orthogonal to c_k then it follows that a_i is not orthogonal to c_k. Applying the difference of means definition we have

$$\bar{y}_i - \bar{y}_{i'} = a_i - a_{i'} + f(b_j \epsilon) , \quad a \perp b \tag{1}$$

$$\bar{y}_j - \bar{y}_{j'} = b_j - b_{j'} + f(c_k \epsilon) , \quad b \perp c \tag{2}$$

and

$$\bar{y}_k - \bar{y}_{k'} = c_k - c_{k'} + f(b_j \epsilon) , \quad b \perp c \tag{3}$$

but b_j is a function of the a_i's, from equation (1). Thus, equation (3) involves the a_i's, and similarly equation (1) will involve the c_k's each indirectly through b_j.

F.5. For a totally orthogonal design the total sum of squares can be partitioned into independent sums of squares for each class of parameters. If the usual normality assumption is applied to the response model then the sums of squares are distributed as independent chi-squares.

F.6. For the "nested hypotheses" definition when σ^2 is known orthogonality implies both independence of hypotheses and independence of the likelihood ratio statistics or equivalently F ratios. Essentially we have hypotheses relating to separate independent experiments, an ideal situation. If σ^2 is unknown, as is generally the case, neither hypotheses nor test statistics are independent. The tests all involve the same denominator or function of it, the error or residual sum of squares. But an orthogonal design still ensures good tests of hypotheses, i.e., good tests of H_1 and H_2 induce a good test of $H_1 \cap H_2$ (see Darroch and Silvey [2] section 6).

F.7. Let D be a block design; then if D is totally orthogonal it is locally connected as defined by Hedayat [6].

5. Conclusion

In this paper orthogonal designs are discussed mainly from the point of view of testing an experimental design to ascertain if it is orthogonal or not. Obviously, the problem of developing simple methods for constructing orthogonal designs is equally important. As yet no specific methods for constructing orthogonal designs have been developed, in fact the only procedure used is to construct the design so as to satisfy whichever definition of orthogonality one prefers.

Non-linear response models pose a problem as to which definition of orthogonality to use. The covariance definition is preferred by the author, since, in general, we are interested in estimation and/or tests of hypotheses of parameters of a design and the covariance definition tells the experimenter what relationships exist between the estimators of different parameters.

References

[1] Yates, F. The principles of orthogonality and confounding in replicated experiments. Jour. of Agric. Sci. 23 (1933) 108-145.

[2] Darroch, J. N. and Silvey, S. D. On testing one or more hypotheses. Ann. Math. Statist. 34 (1963) 555-567.

[3] Seber, G. A. F. Orthogonality in analysis of variance. Ann. Math. Statist. 35 (1964) 705-710.

[4] Scheffé, H. The Analysis of Variance. Wiley (1959).

[5] Graybill, F. An Introduction to Linear Statistical Models, volume 1. McGraw-Hill (1961).

[6] Hedayat, A. A set of conditions which every locally connected design must satisfy. BU-392-M in the Biometrics Unit Mimeo Series, Cornell University.